FA11 10:15

Towards a Methodology for Robust Parameter Identification

Roy S. Smith[†*] and John C. Doyle[*]

† Jet Propulsion Laboratory
California Institute of Technology
Pasadena, CA 91109, U.S.A.

* Electrical Engineering, 116-81
California Institute of Technology
Pasadena, CA 91125, U.S.A.

Abstract

The paper considers the problem of estimating, from experimental data, real parameters for a model with uncertainty in the form of both additive noise and norm bounded perturbations. Such models frequently arise in robust control theory, and a framework is introduced for the consideration of experimental data in robust control analysis problems. If the analysis tools applied include robust stability tests for real parameter variations (real μ), then the framework can be used to address the problem of "robust" parameter identification. While the techniques discussed here can quickly become computationally overwhelming when applied to physical systems and real data, the approach introduces a new way of looking at the identification problem and may be helpful in arriving at a more tractable methodology.

1 Introduction

Developments in robust control theory are providing the engineer with the capability for systematically handling models with increasingly sophisticated uncertainty descriptions, including additive noise together with block structured, norm bounded perturbations. Depending on the assumptions, these perturbations can represent uncertainties arising from "unmodeled dynamics" as well as parametric variations. While there are many important unsolved problems, particularly for models with many perturbations, substantial progress both in theory and computation has made it possible for these techniques to be applied routinely in engineering design.

While the synthesis and analysis theories give a rigorous means of handling a rich class of uncertainty descriptions, the onus is still on the designer to appropriately model the system in this more complicated framework. Robust control design leads to controllers that have guaranteed performance and stability with respect to all members of a model set. Including physically unrealistic models in the model set can make the design conservative as it may be these models which determine the worst case system behavior. The designer therefore wants a model set description which is "tight", in that no physically unrealistic models are included, and yet describes all pertinent behaviors of the physical system.

Identification is the process of generating a model from experimental input-output data. A robust control model is more complicated than the standard linear system transfer function model — the structure of the uncertainty as well as bounds on its size must also be specified. Standard identification techniques for linear models with additive noise are well developed, but currently there are no rigorous means of obtaining a robust control model, except in special cases. The term robust parameter identification is used here to refer to the problem of identifying a model which includes norm bounded perturbations. Note that robust is descriptive of the resulting model — not necessarily the identification technique itself.

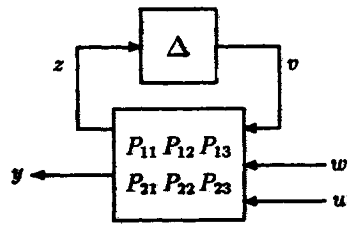

Figure 1: The Generic Structure for Identification and Model Validation Problems

Progress has been made in what is essentially the converse of identification: model validation. The model validation question is "given a robust control model and an experimental input-output datum, does there exist an element of the model set which can exactly describe the observed input-output behavior?" The reader is referred to Smith[1] and Smith and Doyle[2] for details on model validation. The model validation approach sets up a framework for a robustness analysis of systems which include known input and output signals. This framework is used here to consider the robust parameter identification problem.

2 Identification, Model Validation, and Robust Control

Figure 1 shows the structure that will be used throughout as the generic identification and model validation structure. In identification experiments certain inputs to the system are known, so the input is partitioned into u and w with u representing the system inputs that are known, and w represents the unknown inputs from a specified norm-bounded set. The output y represents the measured outputs and is assumed to be known. This model structure, referred to as a linear fractional transformation (LFT), is given by

$$\begin{aligned} e &= \left[P_{21}\Delta(I-P_{11}\Delta)^{-1}P_{12}+P_{22}\right]\begin{bmatrix} w \\ u \end{bmatrix} \qquad (1) \\ &= \mathbf{F}_u(P,\Delta)w \end{aligned}$$

where the Δ is a block structured norm-bounded perturbation. Space constraints preclude an adequate review of these type of models.

An engineer having a physical system and wishing to model it within this framework is immediately faced with a problem: how to first select the system structure and then specify bounds on the perturbations and weights on the input and output sets. An identification methodology is required such that given input-output experiments, and some assumptions on the system, the methodology gives a weighted $\mathbf{F}_u(P,\Delta)$ model which will lead to a satisfactory control design.

2.1 A Black Box Identification Problem

For any given set of data, a large set of models will be able to produce the observed data and the measure of suitability of these will depend strongly on the design performance objectives. To illustrate this, consider the system shown in Figure 2.

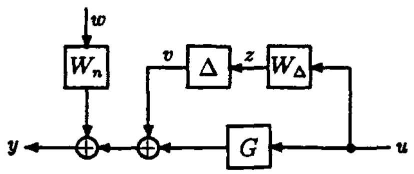

Figure 2: An Example Identification Problem

Given any input-output datum y and u, it is possible to attribute the discrepancies between the nominal behavior ($y = Gu$) and the observed behavior entirely to $W_n w$. Similarly, these residuals can also be attributed entirely to ΔW_Δ. For example, suppose that

$$G = \frac{2}{1+\tau s}, \quad \tau = 10, \quad W_n = 0.1, \quad \text{and} \quad W_\delta = 0.1.$$

and that u, the known input is a sinusoid of unit magnitude at frequency, $\omega = 0.1$ rad/sec. The problem can be considered as a constant matrix problem at $\omega = 0.1$ rad/sec. The input is therefore 1, and the nominal modeled output would be $Gu = (1-j)$. Assume however that y is measured to be $y = (1-1.1j)$ The discrepancy could be attributed entirely to w, so that $w = -j$, $\Delta = 0$, or entirely to Δ ($\Delta = -j$, $w = 0$).

To see the significance of this choice, consider the effect of putting negative feedback from y to u. If $\Delta = 0$ then the resulting closed loop system is stable for any negative feedback gain. Performance measured, say, as the transfer function from w to y in the closed loop system could then be made arbitrarily good. If, on the other hand, $\Delta = -j$, at $\omega = 0.1$, then Δ could be, for example, $\Delta = \frac{(1-10s)}{(1+10s)}$. Note that Δ is stable and that $\|\Delta\| = 1$ at all frequencies. But negative feedback around $G + W_\Delta \Delta$ is unstable if the gain is greater than 10.

In this context the term ambiguity will be used to describe uncertainty about uncertainty. The above system might present little problem in practice as an experiment with $u = 0$ could be used to estimate $W_n w$, and an additional experiment such that $Gu \gg W_n w$ might give a reasonable estimate of ΔW_Δ. While the goal of good experimental design should to reduce ambiguity in the modeling process, it is not practically possible to remove it entirely.

Consider the simple example given above in the context of identification. In a purely "black box" identification scheme G, W_n, and W_Δ are unknown. The problem is even more nonunique than the above example illustrates. One could pick

$$G = \frac{2.21}{1+11s}$$

and account for the observation with $w = 0$ and $\Delta = 0$. The ambiguity could be taken still further — the above example assumed that the Δ block entered the system additively, while it could just as easily be modeled in several other ways.

This discussion illustrates the importance of information about the structure of the model in the identification procedure. First principles modeling will often provide such information and an initial nominal model. Clearly engineering judgement will always be required in the generation of $\mathrm{F}_u(P,\Delta)$ models.

2.2 Model Validation

The model validation theory gives a means of testing a robust control model against the past data. A necessary condition for the suitability of a given model is that it can account for all past data. In the robust control framework this means that for each observed input-output datum there exists a model in the model set able to generate that datum. Consider the data to be a series of experiments. For each experiment the model validation problem is as follows.

Model Validation Problem: *Given a model, $\mathrm{F}_u(P,\Delta)$, and an input-output datum (u,y), does there exist (w,Δ), $w \in \mathbf{BL_2}$, $\Delta \in \mathbf{B\Delta}$, such that*

$$y = \mathrm{F}_u(P,\Delta)\begin{bmatrix} w \\ u \end{bmatrix}.$$

This simply requires that there is an element of the model set and an element of the unknown input signal set such that the observed datum is produced exactly. Any (Δ, w) satisfying this condition will be referred to as admissible. The norm in $\|w\|$ and the norm-bounded set $\mathbf{B\Delta}$ depend on the particular problem.

The model validation test is therefore a necessary condition for any model to describe a physical system. Model validation is a misleading term; strictly speaking, it is never possible to validate a model, only to invalidate it. The fact that every experiment can be accounted for in this manner provides little information about the model and the system. The particular w and Δ do not necessarily bear any relationship to physical signals.

2.3 Solution of the Model Validation Problem

This section will outline the solution of the constant matrix model validation problem. Only a single perturbation block, Δ, is considered for simplicity. The notation used is that of Smith[1]. Here $\|w\|$ will denote the usual Euclidean norm, and $\mathbf{B\Delta}$ will be the set of contractions in the corresponding induced norm.

The assumptions of the model are reformulated as constraints on the signals illustrated in Figure 1. Define $x = \begin{bmatrix} v & w \end{bmatrix}^T$. The model accounting for the data is equivalent to the equality constraint:

$$y - P_{23}u = \begin{bmatrix} P_{21} & P_{22} \end{bmatrix} x. \tag{2}$$

This can be reparametrized by choosing V as a basis for the null space of $\begin{bmatrix} P_{21} & P_{22} \end{bmatrix}$ and x_0 as the solution to Equation 2 that is orthogonal to V. Define the subspace $\mathcal{X}$ as the span of $\begin{bmatrix} x_0 & V \end{bmatrix}$. Then x satisfies Equation 2 if and only if $x \in \mathcal{X}$ and $< x_0, x >= \|x_0\|^2$.

Define a new system, N, by

$$N = \begin{bmatrix} P_{11} \;\; P_{12} \\ \frac{1}{\|x_0\|^2} x_0^* \\ 0 \end{bmatrix} + \frac{1}{\|x_0\|^2}\begin{bmatrix} P_{13} u x_0^* \\ 0 \end{bmatrix},$$

where the zeros are added to make N square. Smith gives a motivation for defining such a system. For notational convenience define two projections, R_1 and R_2, by $R_1x = v$, and $R_2x = w$. Fan and Tits[3] give the following definition of the two block μ problem.

$$\mu = \sup_{\gamma,x} \left\{ \gamma \;\middle|\; \begin{array}{l} \|R_1x\|\gamma \leq \|R_1Nx\| \\ \|R_2x\|\gamma \leq \|R_2Nx\| \end{array} \right\}$$

This formulation leads to an approach for the model validation problem. Introduce the additional constraint that $x \in \mathcal{X}$ into the above problem. Now calculate

$$\gamma_0 = \sup_{\gamma,x\in\mathcal{X}} \left\{ \gamma \;\middle|\; \begin{array}{l} \|R_1x\|\gamma \leq \|R_1Nx\| \\ \|R_2x\|\gamma \leq \|R_2Nx\| \end{array} \right\} \tag{3}$$

Now $\gamma_0 \leq 1$ if and only if there exists an admissible (Δ,w) for the model validation problem. This restricted problem inherits many of the properties of the underlying μ problem. In particular a γ and x can be found that bounds γ_0 above. In the model validation case, the x satisfying the constraints of Equation 3 can be scaled such that $< x_0, x >= \|x_0\|^2$.

The model validation problem for systems decomposes into *independent problems* at each frequency *if one poses the following* optimization problem.

$$\min \|w\| \quad \text{subject to} \quad \|\Delta\| \leq 1 \tag{4}$$
$$\text{and} \qquad y = F_u(N,\Delta)\begin{bmatrix} w \\ u \end{bmatrix}$$

The framework outlined above can treat this problem by redefining γ_0 by

$$\gamma_0 = \sup_{\gamma,x\in\mathcal{X}} \left\{ \gamma \;\middle|\; \begin{array}{l} \|R_1x\| \leq \|R_1Nx\| \\ \|R_2x\|\gamma \leq \|R_2Nx\| \end{array} \right\}.$$

Now γ_0 is the solution to the optimization problem of Equation 4.

The above discussion illustrates that the presence of experimental data, and hence an equality constraint, can be treated by restricting the μ optimization problem to a subspace, $\mathcal{X}$. Note also that the scaling, γ, does not have to apply to all constraints of the problem. The nonhomogeneously scaled problem is referred to as "skewed μ".

The solution to the model validation given by Smith only addresses the case where Δ contains non-repeated full complex perturbation blocks. Similar solution methods are expected to extend this to the repeated complex perturbation case. It will be shown that an extension to the repeated real perturbation case is required to solve the robust parameter identification problem.

3 Robust Parameter Identification

3.1 A Generic Model

The generic identification/model validation structure of Figure 1 is also considered here with one difference. The system is now considered to be fractional on a more general variable, Ψ. Structure is imposed on Ψ by restricting $\Psi \in \boldsymbol{\Psi}$ where

$$\boldsymbol{\Psi} = \mathrm{diag}(\boldsymbol{\Omega}, \boldsymbol{\Theta}, \boldsymbol{\Delta}),$$

where each of the diagonal elements can themselves have block diagonal structure. This represents a division of the fractional parameter into three types. These are:

- $\boldsymbol{\Omega}$ Parameters which have an a priori known relationship to the data and the model. Typical examples would be frequency, or parameters assumed to vary in some a priori known way across the data set, such as when data is taken from similar but nonidentical systems which differ in some way that is modeled exactly. An example of this could be otherwise identical mechanical systems with some differences in the masses of some of the components. Thus in a system model the actual values of $\boldsymbol{\Omega}$ are assumed to be known, and consequently the data can be expressed as a function of Ω: $(y(\Omega), u(\Omega))$.

- $\boldsymbol{\Theta}$ A priori unknown elements which may be fixed with respect to subsets of the elements of $\boldsymbol{\Omega}$. The goal of the identification experiment is to select a suitable $\Theta \in \boldsymbol{\Theta}$. The simplest example of this would be an unknown parameter which was assumed to be fixed across all data sets. A more complicated example would be a parameter which is fixed for the duration of any single experiment, but may vary between experiments. Thus for each experiment, it would be fixed across all values of frequency. In the problems considered here Θ will typically represent real parameter variations.

- $\boldsymbol{\Delta}$ Unknown perturbations which are allowed to vary with respect to all elements of $\boldsymbol{\Omega}$. These might be unknown bounded dynamics where the unknown value is allowed to be different at each frequency.

In setting up a model in this form, one puts as much a priori information as possible into the interconnection structure M. The known correlation between elements of the data, $(y(\Omega),u(\Omega))$, and the model is reflected in the choice of Ω. This framework may seem contrived, but hopefully the significance of these distinctions will be made clear in the examples which follow. The typical robust control model is constructed with $\Psi = \Delta$.

When a robust control model is required for the purposes of design, certain of the elements of Ψ will have known values. In particular $\Omega \in \boldsymbol{\Omega}$ will be completely known. $\Theta \in \boldsymbol{\Theta}$ will also be known, the values having been determined by a prior identification experiment. Δ will be unknown but norm bounded. Similarly, the input w is unknown but norm bounded. The norm bounds on Δ and w can be assumed to be unity without loss of generality.

Closing the loops around Ω and Θ gives rise to the standard robust control model, $N = F_u(M,(\Omega,\Theta))$, with the input-output relationship described by

$$y = F_u(N,\Delta)\begin{bmatrix} w \\ u \end{bmatrix}, \quad \Delta \in \mathbf{B}\boldsymbol{\Delta}, w \in \mathbf{BL_2}.$$

The Ω values are prescribed; finding suitable values for Θ from experimental $(y(\Omega),u(\Omega))$ data is referred to as the robust parameter identification problem and will be considered in the remainder of this paper.

3.2 Ambiguity and the Admissible Set

In an experiment (y,u) is the measured datum, M and Ω, are known by hypothesis, and one wishes to determine Θ such that the assumptions of the resulting robust control model are satisfied. In other words, there exists $\Delta \in \mathbf{B}\boldsymbol{\Delta}$ and $w \in \mathbf{BL_2}$, such that

$$y = F_u(M,(\Psi)\begin{bmatrix} w \\ u \end{bmatrix}, \quad \Psi \in \boldsymbol{\Psi} \tag{5}$$

In general the (Θ,Δ,w) satisfying Equation 5 are not unique reflecting the presence of ambiguity.

The Θ, Δ and w can be traded off, giving different robust control models. Even for a fixed Θ, there will be a set of (Δ,w) satisfying the input-output relationship of Equation 5. This will allow a choice of norm bounds on the Δ and w satisfying the model assumptions, giving a set of possible interconnection structures, N.

Define as the *feasible set*, the set of all (Θ,Δ,w) such that

$$y = F_u(M, \mathrm{diag}(\Omega,\Theta,\Delta))\begin{bmatrix} w \\ u \end{bmatrix}, \quad \Delta \in \mathbf{B\Delta}, \quad w \in \mathbf{BL_2}.$$

Define the *admissible set* as all elements in the feasible set that satisfy the norm bounds: $\|\Delta\| \leq 1$, $\|w\| \leq 1$. These terms are used by Smith[1] for the model validation problem. The extension of the definitions to the more general problem given here is trivial. It should be noted that these sets can be defined equivalently on either (Δ,w), or the signals (v,w).

Ambiguity arises when the admissible set contains more than a single element. The admissible set is defined for a given an input-output experimental datum (y,u) for a particular model, (M,Ω,Θ,Δ). In general ambiguity occurs and one needs a particular method of choosing Θ in order to form the robust control model for design purposes.

3.3 A Performance Function

Consider the problem of selecting Θ from the admissible set to form

$$N = F_u(M, \mathrm{diag}(\Omega,\Theta)).$$

Selecting Θ from the admissible set means choosing a Θ such that there exists Δ and w with (Θ,Δ,w) a member of the admissible set.

The choice of the "best" Θ depends on the intended use of the resulting interconnection structure N. A means of assigning a performance value to Θ is required. Define such a performance function, Φ, where

$$\Phi : (\Omega,\Theta,\Delta,\mathrm{BL_2}) \longrightarrow \mathrm{R}^+.$$

This definition is rather useless; it serves only to formalize the idea of choosing the "best" model for the system. One can then construct an optimization problem for finding a robust control model for the system.

Given M, Ω, $y(\Omega)$, and $u(\Omega)$,

$$\min_{\Delta\in\mathbf{\Delta},w} \Phi(\Omega,\Theta,\Delta,w) \text{ s.t. } (\Theta,\Delta,w) \in \textit{admissible set}.$$

In other words, find the best unknown parameters, Θ, that allow the construction of a model, $F_u(M,(\Omega,\Theta))$, that can account for the observed datum. It is of value to note however that performance functions can be defined which result in the above optimization reducing to a μ calculation of the type discussed in Section 2.3.

Consider the model validation problem in this framework. In this problem a candidate robust control model is compared to the datum. There is no Θ component in Δ. The problem then becomes does there exist an admissible (Δ,w). In the problem studied by Smith[1], the performance function is chosen as

$$\Phi(\Omega,\Delta,w) = \|w\|.$$

Here Ω is a set of frequencies, $\omega_1,\ldots,\omega_n$, and the input and output data are known at each frequency: $y(\omega_i)$, $u(\omega_i)$, $i = 1,\ldots,n$. The choice of $\|w\|$ as the performance function has the advantage that minimizing $w(\omega_i)$, $i = 1,\ldots,n$, minimizes $\Phi(\Omega,\Delta,w)$. The problem then breaks down into n independent problems as the Δ is also independent over frequency: For each ω_i, $i = 1,\ldots,n$, the model validation problem becomes, using the notation $y_i = y(\omega_i)$,

$$\min_{\Delta_i,w_i} \|w_i\| \quad \text{subject to} \quad y_i = F_u(M(j\omega_i),\Delta_i)\begin{bmatrix} w_i \\ u_i \end{bmatrix}, \quad \text{and} \quad \|\Delta_i\| \leq 1.$$

This paper will introduce a framework which can handle certain types of performance functions which do not break down into independent μ type problems for each value of Ω. The robust parameter identification problem has this property.

4 A Series of Examples

The concepts of the previous sections will be illustrated by a series of examples. These will be progressively more difficult problems, leading up to the robust parameter identification problem.

4.1 Calculating a Frequency Response

This example considers the case when $\mathbf{\Psi} = \mathbf{\Omega}$. There are no parameters, Θ, nor is there any uncertainty, Δ. The simplest meaningful problem in such a framework is the calculation of a frequency response.

Consider a known single-input single-output transfer function which is to be calculated at a series of frequencies: ω_i, $i = 1,\ldots,n$.

The example system is described by the transfer function

$$y = Gu = \left[\frac{1}{1+\tau s}\right] u. \tag{6}$$

This can be described as a linear fractional transformation on $1/s$,

$$y = F_u(M, 1/s)u, \quad \text{where } M = \begin{bmatrix} -1/\tau & 1/\tau \\ 1 & 0 \end{bmatrix}.$$

The transfer function can be evaluated for each frequency, ω_i, by calculating the LFT above for $s = j\omega_i$. However all n transfer function values can be calculated simultaneously by duplicating each M_{ij}, $i,j = 1,2$, element of this LFT n times. This gives

$$\hat{y} = F_u(\hat{M}, \frac{1}{s}I_n)\hat{u}, \quad \text{where } \hat{M} = \begin{bmatrix} -1/\tau I_n & 1/\tau I_n \\ I_n & 0 \end{bmatrix}.$$

The notation I_n denotes an identity of dimension n. Note that the dimensions of $\hat{y}$ and $\hat{u}$ are n. Duplicating the original LFT in the above manner gives an LFT fractional on $1/sI_n$. Replace this term by Ω and select $\Omega = \mathrm{diag}(j\omega_1,\ldots,j\omega_n)$. One could alternatively view this operation as making the substitution $s = j\omega_i$, for $i = 1,\ldots,n$ and generating n expressions of the transfer function in LFT form. These transfer functions are then stacked and row and column rearrangement will yield $F_u(\hat{M},\Omega)$ as above.

Now close the loop around Ω, giving

$$N = F_u(\hat{M},\Omega) = \hat{M}_{22} + \hat{M}_{21}\Omega(I_n + \Omega\hat{M}_{11})^{-1}\hat{M}_{12}.$$

The i^{th} diagonal element of N is the transfer function evaluated at $s = j\omega_i$. This method extends trivially to any transfer function by noting that

$$C(sI_{nx} - A)^{-1}B + D = F_u(M, 1/sI_{nx}),$$

where

$$M = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \quad \text{and } A \in \mathbf{R}^{nx \times nx}.$$

The above example considered a SISO system. The discussion applies directly to MIMO systems with the result that N is block diagonal. Clearly, in this example, the problem can be decomposed into n independent problems.

4.2 A Model Validation Example

The interconnection structure for this problem will be formed with $\Psi = \text{diag}(\Omega, \mathbf{\Delta})$. Consider the system shown in Figure 2, where G is that given in the previous section, Equation 6. The weighted model, τ, W_n, and W_Δ in the block diagram, is given. There are measurements of y and u at n frequencies, ω_i, $i = 1, \ldots, n$. The model validation question is simply does there exist $\Delta \in \mathbf{B\Delta}$, and $w \in \mathbf{BL_2}$, such that the input-output observation can be accounted for exactly. Refer to Smith[1] for an experimental method for setting up such a problem.

This can be expressed as an LFT on $1/s$ and Δ.

$$y = F_u(M, \text{diag}(1/s, \Delta)) \begin{bmatrix} w \\ u \end{bmatrix}$$

where

$$M = \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix} = \begin{bmatrix} -1/\tau & 0 & 0 & 1\tau \\ 0 & 0 & 0 & W_\Delta \\ -1 & 1 & W_n & 1 \end{bmatrix}.$$

Note that $M_{11} \in \mathbf{R}^{2\times 2}$. As in the example of the previous section, duplicate this structure n times. Row and column rearrangement and replacing sI_n by Ω, will lead to the following.

$$\hat{y} = F_u(\hat{M}, \text{diag}(\Omega, \Delta_1, \ldots, \Delta_n)) \begin{bmatrix} \hat{w} \\ \hat{u} \end{bmatrix},$$

where $\hat{M}$ is being considered as fractional on a general Ω. Note that $\hat{M}_{11} \in \mathbf{R}^{2n\times 2n}$, $\Omega \in \mathbf{C}^{n\times n}$, and $\hat{w} \in \mathbf{C}^n$. Note also that the duplication of the block structure M has resulted in n perturbation blocks, $\Delta_1, \ldots, \Delta_n$. If Δ is assumed to be independent at each frequency then $\mathbf{\Delta}$ in the $2n \times 2n$ Ψ structure consists of n blocks. If the Δ in the problem was assumed to be fixed for all frequencies the resulting $\mathbf{\Delta} \subset \Psi$ would be a single block repeated n times. This distinction will be made more clearly in the example of Section 4.3.

Again choose a particular Ω, in this case the frequencies corresponding to the input-output measurements, $\Omega = \text{diag}(j\omega_1, \ldots j\omega_n)$. The inputs $\hat{u}$, and outputs, $\hat{y}$, have the following interpretation.

$$\hat{y} = \begin{bmatrix} y(\omega_1) \\ \vdots \\ y(\omega_n) \end{bmatrix}, \quad \hat{u} = \begin{bmatrix} y(\omega_1) \\ \vdots \\ y(\omega_n) \end{bmatrix}.$$

Form N by closing the loop around Ω given above, $N = F_u(\hat{M}, \Omega)$. The resulting $N \in \mathbf{C}^{n\times 2n}$ and now describes the following system.

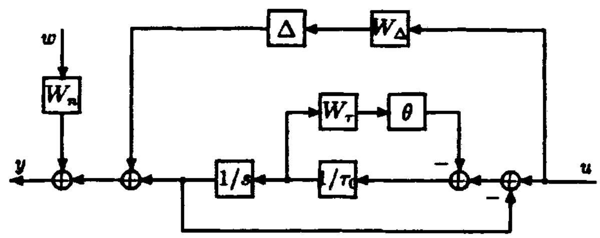

Figure 3: Block Diagram of the Parameter Identification Example

$$\hat{y} = F_u(N, \text{diag}(\Delta_1, \ldots, \Delta_n)) \begin{bmatrix} \hat{w} \\ \hat{u} \end{bmatrix}.$$

The problem is now to test for the existence of n perturbation blocks, Δ, and $w \in \mathbf{C}^n$ such that the model accounts for the data. However the n perturbation blocks, Δ, are in the form of a well defined uncertainty structure. The above problem is therefore a single model validation problem.

One can use the larger problem for N and allow a wider variety of performance functions than that originally considered by Smith. For example, consider

$$\Phi = \|\text{diag}(\Delta, \ldots, \Delta, w(\omega_1), \ldots, w(\omega_n))\|.$$

The optimization problem reduces to a μ problem with the constraint that the signals achieving μ be constrained to a subspace. The optimization problem now has n times as many variables as each of the previous problems. The solution of this problem would give a number, α, such that there exists Δ and w accounting for the observations exactly and at each frequency $\|\Delta\| \le \alpha$, and $\|w(\omega_i)\| \le \alpha$. There exists an admissible (Δ, w) if and only if $\alpha \le 1$.

4.3 A Robust Parameter Identification Example

The example of the previous two sections is extended to a robust parameter identification problem. Consider the time constant, τ, to be uncertain. It is likely that bounds on τ can be postulated for a physical problem. Assume then that τ lies within a range; $\tau_{min} \le \tau \le \tau_{max}$. Reformulate this as

$$\tau = \tau_0 + W_\tau \theta, \quad \text{where } \theta \in \mathbf{R},\ |\theta| \le 1.$$

The nominal, τ_0, and the weight, W_τ are given by

$$W_\tau = (\tau_{max} - \tau_{min})/2, \quad \text{and} \quad \tau_0 = \tau_{min} + W_\tau.$$

Figure 3 shows a block diagram of this system.

This can be reformulated as an LFT on $\Psi = \text{diag}(s, \theta, \Delta)$. Figure 4 illustrates the LFT formulation,

$$y = F_u(M, \text{diag}(s, \theta, \Delta)) \begin{bmatrix} w \\ u \end{bmatrix}. \tag{7}$$

The subblocks of M are:

$$M_{11} = \begin{bmatrix} -1/\tau_0 & -1/\tau_0 & 0 \\ -W_\tau/\tau_0 & -W_\tau/\tau_0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad M_{12} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$

$$M_{13} = \begin{bmatrix} 1/\tau_0 \\ W_\tau/\tau_0 \\ W_\delta \end{bmatrix}, \quad M_{21} = \begin{bmatrix} 1 & 0 & 1 \end{bmatrix},$$

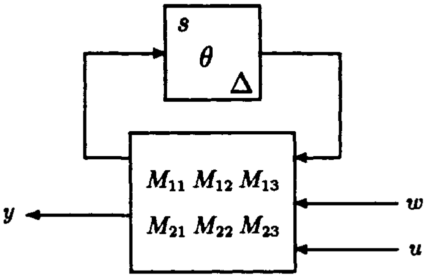

Figure 4: LFT formulation of the Robust Parameter Identification Example

$$M_{22} = \begin{bmatrix} W_n \end{bmatrix}, \quad M_{23} = \begin{bmatrix} 1 \end{bmatrix}.$$

The input-output data is known at a series of frequencies, ω_i, $i = 1, \dots, n$. Again the LFT of Equation 7 is duplicated at each frequency, $s = j\omega_i$, giving the system,

$$\hat{y} = F_u(\hat{M}, \Psi)\begin{bmatrix} \hat{w} \\ \hat{u} \end{bmatrix},$$

where $\Psi \in \mathbf{\Psi} = \mathrm{diag}(\mathbf{\Omega},\mathbf{\Theta},\mathbf{\Delta})$. The subblocks are structured as $\mathbf{\Theta} = \theta I_n$, and $\mathbf{\Delta} = \mathrm{diag}(\Delta_1, \dots, \Delta_n)$. Choose $\Omega = \mathrm{diag}(j\omega_1, \dots, j\omega_n)$, and close the upper loop around Ω, giving $N = F_u(\hat{M}, \Omega)$. The resulting system is

$$\hat{y} = F_u(N, \Psi)\begin{bmatrix} \hat{w} \\ \hat{u} \end{bmatrix},$$

where

$$\Psi \in \mathbf{\Psi} = \mathrm{diag}(\mathbf{\Theta}, \mathbf{\Delta}) = \mathrm{diag}(\theta I_n, \Delta_1, \dots, \Delta_n).$$

As in the example of Section 4.2, the inputs, $\hat{\mathbf{u}}$, and outputs, $\hat{y}$, have the interpretation of vectors over Ω.

Consider the block structure, $\mathbf{\Psi}$, for this system. Δ is included n times as, by the hypotheses of the model, the value of Δ is independent over frequency. However, for the parameter identification problem to make sense, θ is required to be a constant over all frequencies. It therefore enters the structure, $\mathbf{\Psi}$, as a repeated block. Furthermore the value of θ is required to be real. It is the presence of this constant parameter which prevents the problem being decomposed into a smaller independent problem at each of the n frequencies.

The robust parameter identification problem now becomes the problem of finding θ such that there exists $\Delta \in \mathbf{B\Delta}$ and $\hat{w} \in \mathbf{BL_2}$ such that

$$\hat{y} = F_u(N, \mathrm{diag}(\theta I_n, \Delta_1, \dots, \Delta_n))\begin{bmatrix} \hat{w} \\ \hat{u} \end{bmatrix}.$$

The performance function, $\Phi(\theta, \Delta, w)$, allows one to choose the "best" such θ.

The physical requirement that $|\theta| \leq 1$ also allows the choice of performance functions that make the robust parameter identification problem into a model validation problem. As $\|\theta\| \leq 1$, there exists an admissible $(\theta,\Delta,\hat{w})$ if and only if there exists an admissible $(\Psi,\hat{w})$ for the model validation problem where $\Psi \in \mathrm{diag}(\mathbf{\Theta}, \mathbf{\Delta})$. A choice of performance function, Φ, which allows this to be posed as a μ type optimization problem is

$$\Phi(\theta, \Delta, \hat{w}) = \|\mathrm{diag}(\theta I_n, \Delta_1, \dots, \Delta_n, w(\omega_1), \dots, w(\omega_n))\|.$$

Solving this problem by search techniques will yield a θ that is part of an admissible (θ,Δ,w).

5 Solving the Identification Problem

The example of the previous section showed how the robust parameter identification problem can be reformulated as a constant matrix model validation problem in a larger dimension. Consideration of the differences between the model validation problem solved by Smith and the robust parameter identification problem will highlight the areas requiring further research. The major differences are:

- The problem does not decompose as a constant matrix problem at each frequency.
- The model validation problem can only treat nonrepeated blocks with full complex uncertainty. The robust parameter identification problem leads to μ type problems with repeated real uncertainties.
- The model validation solution discussed by Smith considers only one performance function, $\Phi = \|w\|$. Others may be preferred.

The first item poses no difficulty conceptually. The previous discussion has shown that the problem can be posed as a large constant matrix μ type problem. If any parameters are required to be constant across Ω the dimension of the problem can lead to computability problems.

The second item is the most serious difficulty. The upper bound methods for μ extend easily to the repeated complex case. This should also hold for the model validation problems given the strong connection between the μ upper bound methods and Lagrange multipliers for model validation. Extending μ to real valued perturbations is the subject of a great deal of research. See for example, Young[4] and the references contained therein.

The choice of $\Phi = \|w\|$ in the original model validation problem was motivated by the desire to have the problem decompose into independent problems at each frequency. This is no longer an issue in the large interconnection structure — all frequencies are specified as elements of Ω and included in the single problem. More general μ like, performance functions can now be considered.

Acknowledgements

The authors would like to thank Andy Packard, Gary Balas, and Peter Young for their helpful comments. Thanks also to M.

References

[1] R. Smith, *Model Validation for Uncertain Systems.* PhD thesis, California Institute of Technology, 1990.

[2] R. S. Smith and J. Doyle, "Model invalidation — a connection between robust control and identification," in *Proc. Amer. Control Conf.*, pp. 1435–1440, 1989.

[3] M. K. H. Fan and A. L. Tits, "*m*-form numerical range and the computation of the structured singular value," *IEEE Trans. Auto. Control*, vol. AC-33, pp. 284–289, 1988.

[4] P. M. Young and J. C. Doyle, "Computation of the structured singular value with real and complex uncertainties." submitted to 1990 Control Decision Conf., 1990.